SUCCESS STORY

Yahoo! JAPANが、1日100万件のコンテンツ入出稿を担うプラットフォームサービスにNetApp StorageGRIDを採用

コンテンツ入出稿プラットフォーム「Castle」を、コンテナ環境とHTTPアクセスに最適化するためのストレージ最新化。

「未来を創る企業」というスローガンを掲げ、最先端の情報技術で私たちの暮らしをより豊かに、より快適なものに変革し続けるヤフー株式会社。同社が提供する「Yahoo! JAPAN」では、ショッピング、オークション、ニュース、動画、天気・災害、地図など100を超えるサービスやコンテンツが提供されています。この日本最大級のポータルサイトでは、1万社を超える出店者やコンテンツパートナーが「コンテンツ入出稿プラットフォームサービス（Castle）」を通じて毎日およそ100万件のデータの受け渡しを行っています。各種サービスを支えるアプリケーション群のコンテナ化を推進しているヤフーでは、Castleをコンテナ環境とHTTPアクセスに最適化するための見直しに着手。オールフラッシュ構成の高性能オブジェクトストレージアプライアンス「NetApp StorageGRID」を採用し、2020年10月より最新世代Castleの運用を開始しました。

500km
以上離れたデータセンターとリアルタイムでデータを同期

"NetApp StorageGRIDは距離の制約を打ち破るとともに、東西2つのデータセンターを網羅する『単一のストレージ』を扱うイメージで運用できるようになりました"

ヤフー株式会社 サイトオペレーション本部 インフラ技術3部 ストレージ1
三添 匠 氏

データ分析によりYahoo! JAPANのサービスを磨き上げる

国内最大級のポータルサイト「Yahoo! JAPAN」を起点にヤフーが提供する多彩なサービスは、驚くべきスピードで進化しながら新しい価値を創造し続けています。リアル店舗でのショッピングなどの「オフライン」でも、インターネットを活用することで暮らしの利便性は大きく向上できる――ヤフーは、スマホ決済サービス「PayPay」によってそうした新しい体験を私たちの日常にもたらしました。同社 インフラ技術2部DCファシリティの宮田耕一氏は次のように話します。

「Yahoo! JAPAN で2020年11月に実施した『超PayPay祭』では、過去最多のアクセス数を達成しました。このイベントを通じて、オンラインからオフラインまでを網羅するPayPayの利便性を、多くの方に実感していただけたと思います」

2018年10月、ヤフーがソフトバンクと協力してスタートさせた「PayPay」は、わずか2年で3,300万ユーザー/260万加盟店を獲得するなど、日本を代表するキャッシュレスサービスに急成長しました。

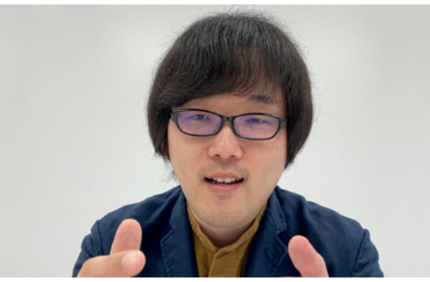
「ヤフーは、お客様に快適なキャッシュレス体験を提供するとともに、ここから得られる膨大な量のデータを分析することでサービスの利便性をさらに磨き上げていきます。こうした取り組みを通じて、お客様にいっそう便利なサービス体験を提供していきます」(宮田氏)

Yahoo! JAPANから提供されるコンテンツやサービスは100を超えており、出店者・コンテンツパートナーとヤフー間では、24時間365日無停止で1日あたりおよそ100万件のデータやファイルがやりとりされています。

「数万社に上る出店者とのデータの受け渡しは、『Castle』と呼ばれるコンテンツ入出稿プラットフォームが一元的に担っています。私たちは、増え続けるコンテンツ入出稿の要求に応えるとともに、コンテナアプリケーションで使い勝手の良いHTTPアクセスに対応するために、新世代Castleの構築に着手しました。新たに採用したのはオブジェクトストレージ『NetApp StorageGRID』です」(三添氏)

高性能オブジェクトストレージアプライアンス「NetApp StorageGRID」を採用

NetApp StorageGRIDは、世界中で豊富な実績を誇るオブジェクトストレージ製品(アプライアンスまたはソフトウェア)です。Amazon S3互換のオブジェクトストレージであり、最大560PBのデータを最大16か所のデータセンターに分散配置して「単一のネームスペース」で管理できるスケーラビリティを備えています。



ヤフー株式会社 サイトオペレーション本部
インフラ技術3部 ストレージ1
三添 匠 氏

「ヤフーでは、アプリケーションのコンテナ化・マイクロサービス化を推進しています。インフラ側もコンテナアプリへの最適化を段階的に進めており、アプリケーションから利用しやすいHTTPアクセス可能なオブジェクトストレージへの需要が高まってきています」と同社 サイトオペレーション本部 インフラ技術3部 ストレージ1の三添匠氏は話します。

Castleでは、数KBから数GBまで画像や動画など、様々なサイズ・形式のデータやファイルを扱います。Castleがデータを受領すると、即時でYahoo! JAPANのサービスに反映され、ショップの商品や在庫情報、ニュースや動画サービスなどが更新されていく仕組みが構築されています。

「Castleは、出店者・コンテンツプロバイダーとヤフーがデータ入出稿を行う唯一の窓口です。天気や災害、交通情報など、暮らしの安心と安全に欠かせない情報も扱うため、サービス停止やデータ喪失は許されません。私たちは新世代Castle構築に際して、サービス継続とデータ保護のための仕組みを強化する方針を掲げました。具体的には、500km以上離れたデータセンターとリアルタイムでデータを同期させ、一方に問題が発生したときでもCastleのサービスを継続できるようにすることを必須の要件としました」（宮田氏）

ヤフーが新世代Castleのストレージに求めた要件を整理すると次の通りです。

①Amazon S3互換でHTTPアクセス可能なオブジェクトストレージであること
②500Km以上離れたデータセンター間とリアルタイムでの完全同期・サービス継続が可能なこと
③データが入稿された際に即時に後続のシステムに通知が可能なこと
④複数のサービスに対するマルチテナント管理をAPIを介して行えること
⑤処理要求の増大に際して、容量と性能を容易に拡張可能なこと
⑥一定期間経過後に不要になった入稿データを自動で削除しスペースを最大限活用できること

「NetApp StorageGRIDはすべての要件を満たしており、中でもオールフラッシュ構成の高性能アプライアンス（SGF6024）を導入できることが大きな魅力でした。また、距離の制約なく2つのデータセンター間で巨大なネームスペースを構成できること、マルチテナント環境で柔軟にデータ保護ポリシーを適用できるメリットも高く評価しました」（三添氏）

ヤフーでは数100に及ぶ検証項目を用意し、およそ半年をかけた入念な実機テストによりNetApp StorageGRIDの性能や機能を評価。Amazon S3プロトコルへの準拠と、拠点間の同期を確実に行えることが確認されました。

「稼働中のNetApp StorageGRIDのケーブルやコンポーネントを引き抜いても、確実にシステムを切り替えてサービスを継続できることを確認しました。システムの可用性・耐障害性の面で優れていると判断しています」（宮田氏）

「データハブ」への活用を可能にするオールフラッシュ構成

新世代Castleの中核に導入されたNetApp StorageGRID SGF6024は、オールフラッシュ構成の高性能オブジェクトストレージアプライアンスです。一般に、オブジェクトストレージは大容量のデータを安価に管理する目的で使われますが、ヤフーの狙いはこれとは大きく異なるものでした。

「Castleは『データハブ』としてすべてのデータ入出稿を担う唯一の窓口です。NetApp StorageGRID SGF6024は、オールフラッシュならではの高い性能を活かして、膨大な件数のデータ入出稿を高速かつ安定的に処理します。オブジェクトストレージでありながら、リアルタイム性の求められるトランザクションにしっかりと応えてくれています」（三添氏）

新世代Castleでは、500km以上離れたデータセンターに設置されたNetApp StorageGRID間でリアルタイムでのデータ二重化を実現。ILMでオブジェクトを2つのデータセンター間で最適に配置するとともに、性能と容量効率をバランスよく両立させています。

ヤフー株式会社 サイトオペレーション本部
インフラ技術2部 DCファシリティ
宮田 耕一 氏

「既存のストレージ環境はリアルタイムデータ同期に距離制限があり、首都圏と関西圏でBCP/DRシステムを構築することは困難でした。NetApp StorageGRIDは距離の制約を打ち破るとともに、東西2つのデータセンターを網羅する『単一のストレージ』を扱うイメージで運用できるようになりました」と三添氏は評価します。

サービス単位で柔軟にデータ保護ポリシーを適用

宮田氏は、「オブジェクトやメタデータを利用してデータ保護ポリシーを適用できる機能に注目した」と話します。NetApp StorageGRIDの独自機能である『ILMポリシー』は、テナントやバケットごとに設定した条件に基づき、オブジェクトの保存場所、保護方法、保護期間などを最適化できます。

「ILMポリシーを使ってライフサイクルを設定し、期限が来た時点でオブジェクトを自動的に削除するような運用が可能になりました。また、ノード障害や広域災害時でもサービスを継続できるよう、ノード間・データセンター間でデータ配置をきめ細かく設定できました。ILMポリシーは、私たちにとってまだまだ幅広く応用できる可能性

がありそうです」（宮田氏）

さらに、三添氏・宮田氏らは、新世代Castleのための運用監視環境を自作しました。

「標準的な監視項目に加えて、中長期の視点でアクセス数や性能の傾向を把握できるよう工夫しました。NetApp StorageGRIDがPrometheusを搭載しており、必要なメトリクスを取得しているおかげで、Grafanaを利用した監視用システムを短期間で実装できました」（宮田氏）

HTTPアクセス可能な高性能ストレージ環境へのニーズ

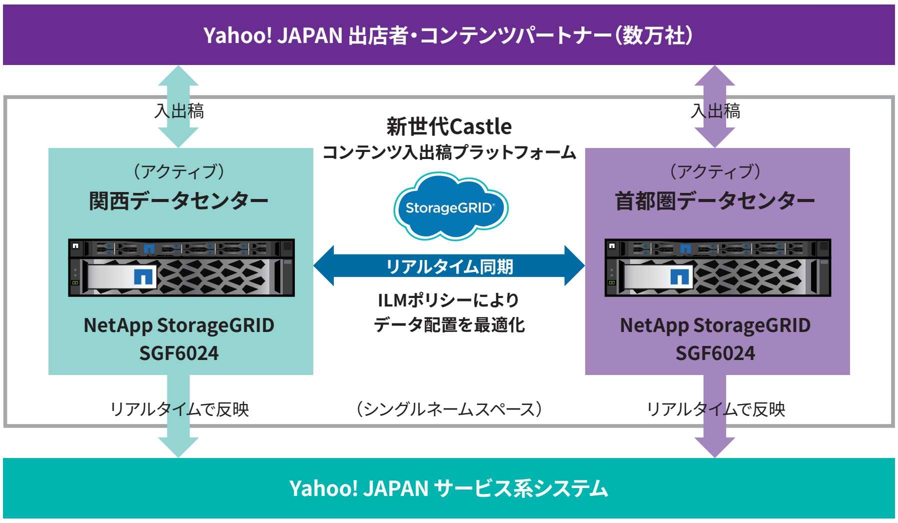
「ILMポリシーを使える高性能オブジェクトストレージアプライアンス」というNetApp StorageGRIDの独創性が、新世代Castleを新しいステージへと引き上げました。

「HTTPアクセス可能な高性能ストレージ環境は、アプリケーションのコンテナ化が進む過程でさらにニーズが高まってくるでしょう。ILMポリシーを使えるNetApp StorageGRIDの応用範囲は広いはずです」と三添氏は期待を示します。さらに宮田氏が次のように続けます。

「リアルタイムデータ同期の距離の制約がなくなり、ヤフーが利用する国内のデータセンター全域を網羅できるようになったことは、ビジネス視点から見ても大きな成果ですね。もしデータセンターを変更することになったとしても、NetApp StorageGRIDならデータ移動に悩む必要はありません」

三添氏、宮田氏は次のように話して締めくくりました。

「ILMポリシーを使ったデータの最適配置を実装する過程では、ネットアップの技術支援を受けながら確実に課題を解決していくことができました。米本社の開発チームを巻き込んだやりとりは実にスムーズで、物理的な距離を感じさせず、私たちと一体となってプロジェクトに取り組んでくれたネットアップには本当に感謝しています。今後も最新のテクノロジーで私たちのチャレンジを支えてもらえることを期待します」



NetApp products

StorageGRID

Protocols

Simple Storage Service

ネットアップ合同会社
TEL：03-6870-7600
Email：ng-sales-inquiry@netapp.com

ネットアップは、ハイブリッド クラウドのデータに関するオーソリティです。クラウド環境からオンプレミス環境にわたるアプリケーションとデータの管理を簡易化し、デジタル変革を加速する包括的なハイブリッド クラウド データサービスを提供しています。グローバル企業がデータのポテンシャルを最大限に引き出し、お客様とのコンタクトの強化、イノベーションの促進、業務の最適化を図れるよう、パートナー様とともに取り組んでいます。

詳細については、www.netapp.com/jaをご覧ください。

NetApp

Google Cloud、Google Compute Engine および Google Cloud Storage は Google LLC の商標です。
© 2021 NetApp, Inc. All rights reserved. 記載事項は、予告なく変更される場合があります。内容の一部または全部をNetApp, Incの許可なく使用・複製することはできません。NetApp、NetAppロゴ、SolidFireは、米国およびその他の国におけるNetApp, Incの登録商標です。その他記載のブランド・製品名は、それぞれの会社の商標または登録商標です。CSS-7178-0421